European Journal of Agronomy 150 (2023) 126930

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Soil properties zoning of agricultural fields based on a climate-driven spatial clustering of remote sensing time series data

Francesco Reyes [a,b], Raffaele Casa [a,*], Massimo Tolomio [a,c], Michele Dalponte [d], Nada Mzid [a]

[a] DAFNE, Università della Tuscia, Via San Camillo de Lellis, 01100 Viterbo, Italy
[b] Dipartimento di Scienze della Vita, Centro Biogest-Siteia, Università degli Studi di Modena e Reggio Emilia, Via Amendola 2, 42122, Reggio Emilia, Italy
[c] Wageningen University, Plant Production Systems, Radix Nova (building 109), Bornsesteeg 48, 6708PE Wageningen, the Netherlands
[d] Research and Innovation Centre, Fondazione E. Mach, Via E. Mach 1, 38098 San Michele all'Adige, TN, Italy

ARTICLE INFO

Keywords:
Clustering
Microclimate
SPEI
Soil properties
Management zones

ABSTRACT

The identification of zones within an agricultural field that respond differently to environmental factors and agronomic management is a key requirement for the adoption of more precise and sustainable agricultural practices. Several approaches based on spatial clustering methods applied to different data sources, e.g. yield maps, proximal sensors and soil surveys, have been proposed in the last decades. The current availability of a huge amount of free remote sensing data allows to apply these approaches to agricultural areas where ground or proximal data are not available. However, in order to provide useful agronomic management information, it is essential that the zoning obtained by clustering is linked to the underlying spatial variability of soil properties. In this work we explore the hypothesis that the response of crop vigor to temporal climate variability, assessed by remote sensing data time series, selected to correspond to specific growth phases and seasonal climate patterns, provides indications on the variability of soil properties within agricultural fields, for both herbaceous and tree crops. NDVI time-series for 38 years (1984–2021) were obtained for fourteen non-irrigated herbaceous and tree crop fields in Central Italy, from multispectral satellites data (Landsat 5/7/8, Sentinel 2). The Standardized Precipitation-Evapotranspiration Index (SPEI) was used to classify time series into three climatic classes (dry/normal/wet) for five different periods of the growth season, covering the main phenological phases. K-means clustering was used to identify patterns of crop growth from climatically classified image sets, as well as for all the bulked images for comparison (bulk clustering). Clustering results were compared with soil maps obtained from spatialized ground data, for soil texture (clay, silt and sand), soil organic matter and available soil water (ASW). The agreement between the different clustering results and soil maps was assessed by the Adjusted Rand Index. Agreement with soil maps varied depending on the field, the phenological phase considered and the soil property considered. Climate driven clustering from long, late growth season periods best matched soil properties, both for herbaceous and tree crops, despite being based on a limited number of images. The clustering from images spanning a longer growth period for dry years systematically outpaced the bulk clustering for silt, sand and ASW, while the clustering for normal climatic conditions was the best for organic matter. The performance of the matching between clustering and soil maps increased with soil variability significantly more ($P < 0.05$) than in the bulk clustering (mean slopes respectively 0.468 ± 0.167; 0.113 ± 0.270). The integration of the SPEI climatic index into the clustering procedure systematically improved the identification of zones with homogeneous soil properties, highlighting that a greater attention should be posed to the climate-crop-field interactions when using remotely sensed images.

1. Introduction

The ability to quantify the patterns of soil spatial variability within agricultural fields accurately, rapidly and at low cost is a key requirement for the implementation of advanced farm management practices to optimize the use of limited resources such as agricultural inputs (water, energy, and fertilizer), while maximizing yield (Castrignanò et al., 2018; Maestrini and Basso, 2018).

* Corresponding author.
E-mail address: rcasa@unitus.it (R. Casa).

https://doi.org/10.1016/j.eja.2023.126930
Received 6 June 2022; Received in revised form 11 July 2023; Accepted 30 July 2023
Available online 10 August 2023

In the context of precision agriculture, partitioning agricultural fields into relatively uniform management zones (MZs) is a common approach for addressing spatial variability, that allows to apply agricultural inputs depending on site-specific soil and crop conditions (Mulla, 2013; Zhang et al., 2002). MZs are sub-units of a field, expected to maximize homogeneity in crop response to environmental and agronomic factors within the sub-unit, while maximizing heterogeneity among them. Homogeneous applications of agricultural inputs to each MZ have been shown to suit crop requirements more efficiently than when applying a uniform dose to the whole field (Ge et al., 2011). Precision management through MZs has been shown to increase crop yield (Breunig et al., 2020), reducing at the same time the environmental impact associated, for example, to fertilizer or herbicide applications (Gavioli et al., 2019), as compared to conventional uniform field management. Several approaches have been proposed to characterize soil spatial pattern, as well as agronomic and yield-limiting factors (Nawar et al., 2017), most of which are based on the exploitation of dense and multivariate information layers (Dhawale et al., 2014; Saifuzzaman et al., 2019). Clustering methods have been proposed to include information layers regarding the soil (Ortega and Santibáñez, 2007), topographic maps (Johnson et al., 2003), yield maps, data from several seasons (Diker et al., 2004) or a combination of these methods (Franzen and Nanna, 2002).

The determination of MZs through clustering approaches based on remote sensing (RS) data has become increasingly popular (Albornoz et al., 2018; Castrignanò et al., 2017; Dhawale et al., 2014; Saifuzzaman et al., 2019), thanks to the ease of access of a large amount of up-to-date spatial RS information regarding agricultural fields. RS of vegetation alone, or in combination with other information layers, such as proximal sensing data, has been widely used as stand-alone information to define MZs. Breunig et al. (2020) defined MZs starting from aboveground biomass (AGB) estimated via remote sensing. Maestrini and Basso (2018) suggested combining RS-based crop information (red-band spectral reflectance, NDVI and plant surface temperature) with the use of historical yield maps from multiple years. Their results showed that the best predictor of within-field spatial variability is the historical yield map for areas with stable productivity (high, medium, low) over time, while spatial variability of areas with unstable productivity was better predicted by within-season RS images, thus creating an operational basis for the definition of MZs. Fontanet et al. (2020) delineated MZs based on crop (remotely sensed NDVI time-series) and soil (soil moisture sensor measurements) information to support management decisions that considered the interactions between crop production, environment, and agronomic operations. Ohana-Levi et al. (2019) and Saifuzzaman et al. (2019) combined topographic data, soil information (electrical conductivity and bare soil indices) and crop information (satellite vegetation indices) to determine MZs.

Understanding the drivers of the spatial distribution of yield within the field is essential to define uniform management zones (Schepers et al., 2004). However, in-season measurements are not sufficient to fully understand the complex interaction of soil characteristics, weather and landscape position that concur in determining crop yield, especially in unstable areas (Maestrini and Basso, 2018). Particularly, the weather pattern plays an important role in determining yield (Chen et al., 2004). Temperature and precipitation patterns during the growing season are capable of explaining a great deal of crop yield variability (Moore and Lobell, 2014; Osborne and Wheeler, 2013; Powell and Reinhard, 2015; Ray et al., 2015). Relationships between weather and yield depend on the interactions between local biophysical conditions and crop management practices (Delerce et al., 2016). For example, Basso et al. (2011) insisted that spatial variability per se, estimated during a single season, is not sufficient to plan fertilization, if weather interactions are not considered. In proposing an optimized strategy for wheat fertilization in Southern Italy (Mediterranean climate), the authors demonstrated that high yielding zones exposed to low precipitation rates may face a reduction in grain yield in case of high nitrogen fertilization. This is due to the fact that high nitrogen doses in very vigorous crops may cause the use of relatively high water volumes before flowering stage, thus potentially limiting water availability in following phenological phases, with consequent premature crops senescence and low grain yield.

Information relevant for delineating sub-units of a field is easy to obtain at the required spatial resolution, from free access remote sensing data, e.g. Sentinel-2 and Landsat-8, in terms of normalized difference vegetation index (NDVI). NDVI is related to green biomass and thus to chlorophyll content and leaf area index (LAI). Therefore, unexpectedly low values occurring within part or the whole field may result from nitrogen deficiency, but also from other plant stresses, such as water deficit (Wang et al., 2016). Despite being easier to obtain than ground data, the limitations of using crop vigor indices such as NDVI, to identify stable MZs within cropped fields, also persist when multi-temporal and multi-annual time series are considered, and integration with ground data is still advised (Marino and Alvino, 2018).

Climatic indices like the Standardized Precipitation Index (SPI) (Mckee et al., 1993) and the Standardized Precipitation-Evapotranspiration index (SPEI) (Vicente-Serrano and Beguería, 2018) can be used to classify a cropping season into climate classes. Both indices provide a functional definition of drought. SPI and SPEI use long-term weather data (at least 30–40 years) of a pre-specified period of the year to quantify, for each year, how the selected period is far from the normalized long-term climate average of that same period over multiple years, in terms of wetness or dryness. Therefore, they may prove useful to categorize NDVI datasets depending on the meteorological conditions affecting crop growth, at the time the original multispectral images were acquired.

In this paper we propose to classify the different parts of each growing season occurring for an agricultural field based on a climatic index (SPEI). These are used to discriminate years with similar weather pattern (dry, normal, wet) and aggregate the corresponding imagery on which to perform multi-temporal clustering of RS-derived NDVI maps, as per climate category. This climate-driven clustering is proposed to delineate MZs that should be representative of crop vigor, accounting for the variation brought in by meteorological conditions. Since crop vigor for a given climatic condition depends largely on soil properties, it can be used to map field-scale soil heterogeneity. We compared clustering results to ground-truth maps of soil texture and available soil water content. The agreement between the NDVI clustering of climatically classified images and soil maps is quantified and compared to the agreement achieved by a more classical multi-temporal NDVI bulk clustering, that does not differentiate growth seasons in terms of weather pattern. The objective was thus to investigate if the proposed climate-driven clustering could be used to delineate field zones based on the actual spatial variability of soil properties of agronomic relevance.

2. Materials and methods

For this study, 14 agricultural fields were used, 12 with herbaceous crop rotations and 2 with trees crops, distributed across two Central Italian regions: Umbria and Lazio (Fig. 1). Of the fields 5 were in hilly areas and were sloping or gently sloping whereas the others were in flat terrain (Table 1). Multi-temporal datasets of images acquired by Sentinel-2 and Landsat satellites acquired during 37 years (1984–2021) were gathered and used to compute the Normalized Difference Vegetation Index (NDVI) time series for these fields (details are provided in Section 2.1) (Fig. 2). The NDVI time series were pre-processed to improve consistency and remove potential outliers (see Section 2.1). Images were filtered to cover the growing season only of non-irrigated herbaceous and tree crops (see Section 2.1.1), in order to exclude the influence of irrigation on temporal and spatial patterns of crop vigor. Meteorological data acquired nearby the monitored fields were collected and gap filled (see Section 2.1.2). They were then used to compute the Standardize Precipitation Evapotranspiration Index (SPEI)

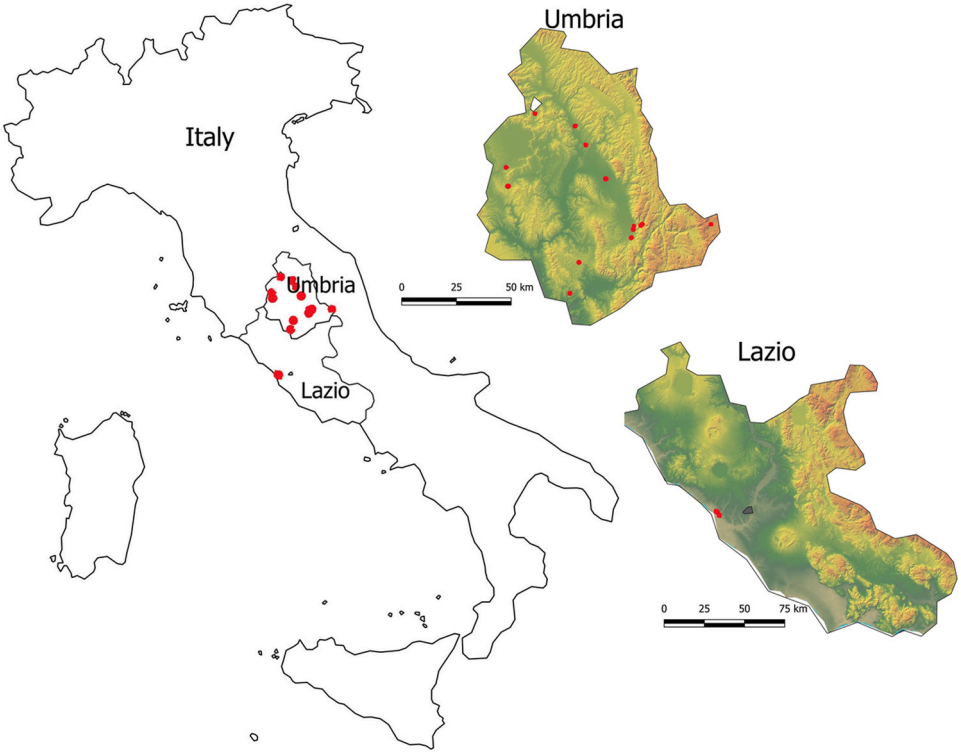

Fig. 1. The study areas containing the fields in the Umbria and Lazio regions. For details of the locations and characteristics of the fields see Table 1.

Table 1
List of agricultural fields used in the present study. Environmental covariates employed to produce soil maps for each field are indicated by x. The spatialization mapping procedure is also reported: BK: Block Kriging, MLR: Multiple Linear Regression, RK: Regression Kriging.

Field name	Region	Location and geographic coordinates	Area (ha)	Topography	Crop type	Covariates				Mapping procedure
						Soil sampling points	Apparent soil resistivity	DEM	BSI	
B030	Lazio	Maccarese (lat. 41.895° N, lon. 12.208° E)	17.53	flat	herbaceous	100	-	-	x	BK
B041		Maccarese (lat. 41.891° N, lon. 12.217° E)	26.76	flat	herbaceous	97	-	-	x	BK
B064		Maccarese (lat. 41.879° N, lon. 12.230° E)	12.17	flat	herbaceous	97	-	-	x	BK
B071		Maccarese (lat. 41.087° N, lon. 12.234° E)	26.59	flat	herbaceous	93	-	-	x	BK
Biavati	Umbria	Paciano (lat. 43.034° N, lon. 12.049° E)	5.72	sloping	herbaceous	4	x	x	x	MLR
Rio Grande		Pieve Pagliaccia (lat. 43.137° N, lon. 12.488° E)	6.45	sloping	herbaceous	5	x	x	x	MLR
Tamburini		Montecastrilli (lat. 43.654° N, lon. 12.472° E)	3.33	sloping	herbaceous	5	x	x	x	MLR
Appoloni		Campello sul Clitunno (lat. 42.809° N, lon. 12.772° E)	2.80	gently sloping	herbaceous	5	x	-	x	MLR
Bachetoni		Campello sul Clitunno (lat. 42.814° N, lon. 12.812° E)	7.51	gently sloping	herbaceous	5	x	-	x	MLR
Bennicelli		Solfagnano (lat. 43.214° N, lon. 12.427° E)	5.07	sloping	vineyard	18	x	x	-	RK
Bertoldo		Città della Pieve (lat. 42.957° N, lon. 12.062° E)	6.66	sloping	herbaceous	24	x	x	x	RK
Ciri		Cannara (lat. 43.000° N, lon. 12.607° E)	9.95	flat	herbaceous	25	x	-	x	RK
Del Rio		Campello sul Clitunno (lat. 42.795° N, lon. 12.767° E)	9.38	flat	herbaceous	5	x	-	x	MLR
Zuccari		Cortaccione (lat. 42.760° N, lon. 12.759° E)	2.95	flat	olives	5	x	-	x	MLR

(Vicente-Serrano and Beguería, 2018) for periods of different duration, i.e., three and five consecutive months, in order to classify growing seasons into three climatic classes: dry, normal and wet. Satellite images covering specific phenological phases of the crops were thus classified based on the SPEI (see Section 2.1.2). Within-field patterns of crop growth were identified by applying a k-means clustering to images of the same climatic class. These patterns were compared to those detected by a conventional bulk clustering, without considering any climatic

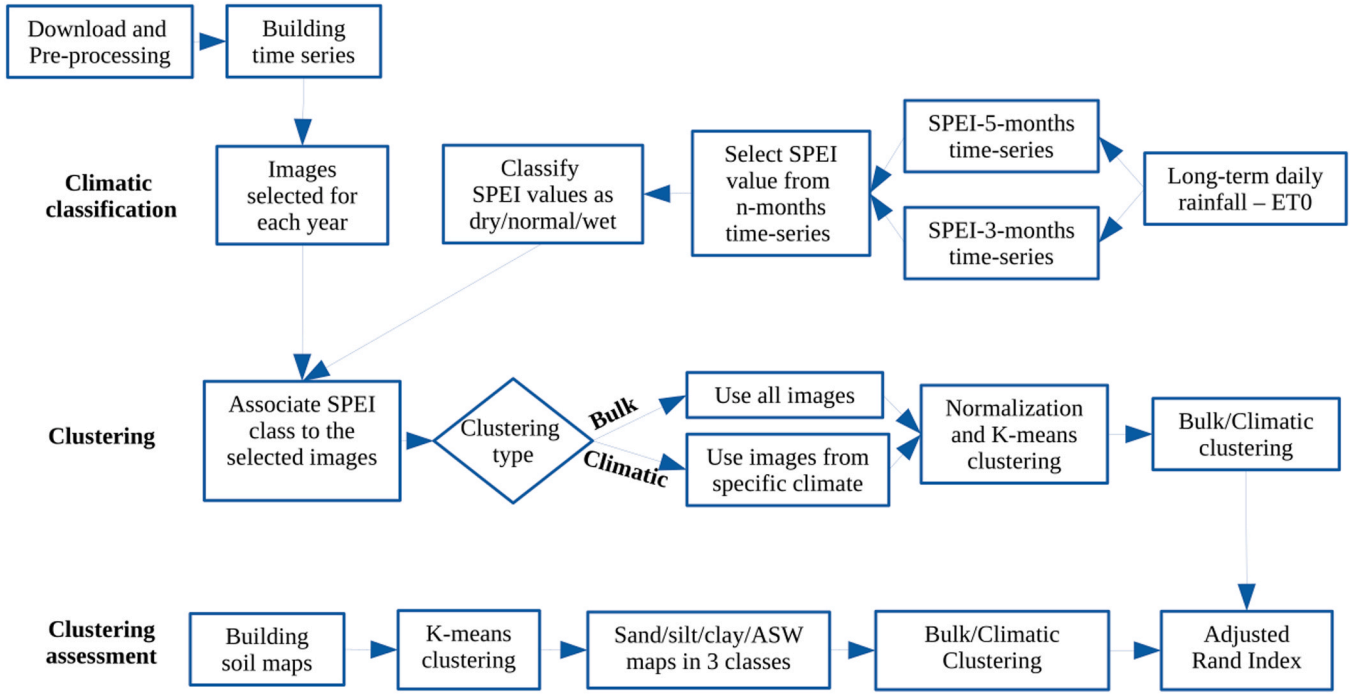

Fig. 2. Overall data processing workflow.

classification of the images (see Section 2.1.3). Soil texture, soil organic matter and available water content maps were obtained from a field survey, followed by spatial analysis and application of pedotransfer functions (see Section 2.2). The agreement between spatial patterns of plant growth and soil properties was assessed by means of a spatial index (see Section 2.2.1).

2.1. Satellite and meteorological data processing

In order to depict crop responses for the same fields under different weather conditions occurring during the crop growth season, we used long term, free access imagery, acquired by Landsat 5, 7 and 8 (30 m spatial resolution) and Sentinel-2 (10 m resolution) satellites. Polygon vector files of the agricultural fields borders, extended by a 60 m buffer were used, within the Google Earth Engine platform (GEE) (Gorelick et al., 2017), to identify images available for the areas of interest. Here, all clear sky images covering the fields and acquired by the satellites during the period January 1984 - June 2021 were considered (Table 2). The NDVI was computed for each image on the area covering the field. Bands 8 (red) and 12 (NIR) were used for Sentinel-2 (S2); bands 3 and 4 for Landsat-5 (L5) and for Landsat-7 (L7); bands 4 and 5 for Landsat-8 (L8). The resulting field-specific NDVI time-series were downloaded from the GEE platform. All following processing was performed in the R statistical environment (R Core Team, 2020).

When multiple images were available for the same field and date only the image with the highest spatial resolution and acquired with the most recent satellite was kept.

Table 2
Details on satellite imagery used in this study.

Satellite	Starting date	Ending date	Filter applied to exclude non-clear sky images
Sentinel-2	3/28/2017	6/30/2021	Image is excluded if any pixel is classified as cloud or cirrus in band QA60
Landsat-8	4/11/2013	6/30/2021	Image is excluded if any pixel is classified as cloud or shadows in band pixel_qa
Landsat-7	1/1/1999	6/30/2021	
Landsat-5	1/1/1984	5/5/2012	

The signal from pixels lying along the edges of agricultural fields is likely affected by elements other than the crop (such as roads, buildings, dikes, or other vegetation types). To avoid their inclusion in the clustering algorithm, images were inspected to exclude potential outliers near the borders of the fields. For any image, pixels were split in two groups: border pixels lying near the field border (less or equal to 30 m for Landsat; less or equal to 10 m for S2), and inner pixels. Border pixels whose values exceeded 1.5 times the Inter Quartile Range (IQR) of NDVI values of the inner field pixels, were considered outliers and were assigned the NA value. A moving window (3 $\times$3 pixels) was then used to iteratively fill the NAs values with the median NDVI value of non-NA pixels in the window. This improved the correspondence between clustering and soil maps (Section 2.2.1) as compared to clustering in the absence of this pre-processing step.

To ensure co-registration, images were then resampled (bilinear method) to a 5 m resolution (Pascucci et al., 2018), reprojected to the coordinate reference system of the corresponding soil map, and finally cropped to its extent.

2.1.1. Assembly of NDVI time-series

To consider the influence of climate on crop response, while reducing the influence of management practices such as irrigation, the images were only selected from non-irrigated fields. Considering herbaceous crops, reliable data on irrigation practices for long time periods were not available. For this reason, we decided to use only images from winter crops, as these are usually not irrigated in Central Italy. For any growth season, images acquired between the previous November 1st and September 31st were selected. The median NDVI value for the field was calculated for each day. A LOESS function was fit to the yearly time-series of median NDVI values computed from each image. Local maxima and minima of the fitted function were extracted. To select winter crops, only data with NDVI maxima occurring between December 1st and May 30th were kept. The local minima immediately preceding and following the selected maxima were considered as the earliest and latest images representing the winter crop for the current year. Other images of the same year were excluded from further processing.

Regarding tree crops, changes in crop type are likely to occur over several decades, related to tree aging and replanting. In this study, crop type changes occurred in both the considered fields during the last 20 years. In addition, one of the fields was equipped with an irrigation system in recent years (farmer's communication), time from which images do not comply any more with the assumptions of the study. This implied that the NDVI time series portraying tree fields were reduced from 40 years each, to just nine and 12 years. The images considered were those acquired between April 1st and September 30th or October 30th, respectively for grapevine and olive, as approximately corresponding to their vegetative periods.

Finally, independent of the crop type, only images containing more than 70% of pixels with NDVI equal or higher than 0.7 were considered as representing fields covered by the crop and thus kept for further analysis (Pascucci et al., 2018).

2.1.2. Meteorological data processing and SPEI calculation

Weather data were collected from the historical dataset of the hydrographic service of Umbria Region and the Regional Agency for Development and Innovation in Agriculture of the Lazio Region (ARSIAL - L'Agenzia Regionale per lo Sviluppo e l'Innovazione dell'Agricoltura del Lazio). Daily rainfall, minimum and maximum temperatures were collected from one or two of the closest weather stations to each field, within a maximum range of 5 km, depending on data availability. Some stations were equipped with only a rain gauge. Since meteorological monitoring was discontinuous, the Joint Research Center (JRC) dataset Agri4Cast Gridded Agro-Meteorological Data in Europe version 3.0 (https://agri4cast.jrc.ec.europa.eu/DataPortal/) was used to fill missing values. This dataset provides daily meteorological data spatialized from weather stations observations on a 25 × 25 km grid between 1979 and 2020. Observed (stations) and predicted (JRC) meteorological variables were well correlated, respectively with a R^2 of 0.64 and 0.61 for 3-months and 5-months cumulative rainfall, and with a R^2 of 0.97 and 0.90 for daily minimum and maximum temperature. This provided a satisfactory fit for the purpose of gap filling, considering how these variables and timeframes were used for calculating a climatic index, as detailed in Section 2.1.2. This permitted us to obtain a weather dataset of at least 41 years (1979 – 2021) for each field.

Daily reference evapotranspiration (ETo) was then calculated from minimum and maximum temperature and extraterrestrial radiation obtained from the station latitude, using the Hargreaves – Samani equation (Danlu et al., 2020).

The climate characterizing each field was synthesized by means of the SPEI climatic index (Vicente-Serrano and Beguería, 2018), for periods of 3 and 5 months covering specific phenological phases, as detailed hereafter. The SPEI index (Standardized Precipitation Evapotranspiration Index) is calculated from rainfall and reference evapotranspiration of the considered period, comparing them to their long-term (> 30 years) values, resulting in a functional definition of drought. Calculations were performed using the "SPEI" R package (Beguería and Vicente-Serrano, 2018), obtaining a 3-months and a 5-months SPEI time-series for each field.

To climatically characterize different phases of crop development in a field, we considered SPEI values for 3 shorter and 2 longer crop growth periods, using the 3-months and 5-months SPEI time-series respectively.

For herbaceous crops, the short periods roughly correspond to the following developmental phases: a) the period around crop emergence, i.e. from November to January included (coded as SW1h, i.e. for short-winter-herbaceous); b) the maximum LAI development period, i.e. from January to March (SW2h); and c) the period around flowering and grain filling, i.e. from March to May (SSh, short-spring). The longer periods correspond to longer parts of the crop growth season: the five months from November to March (LW, long-winter) and from January to May (LSh, long-spring-herbaceous).

For tree crops, the short periods correspond to: a) budburst from January to March (SWt); b) early leaf development, i.e. from March to May (SSt); and c) the warmest summer period during yield formation, from June to August (SSmt, short-summer). Long periods for tree crops were: a) from release of dormancy to leaf development, i.e. from January to May (LSt1) and b) most of the vegetative period, i.e. from April to August (LSt2).

SPEI time-series values for the above periods were then classified as corresponding to: dry periods with SPEI less than − 1; normal periods with SPEI comprised between − 1 and 1; wet periods with SPEI higher than 1. Such a classification was preferred to the usual eight classes employed for SPEI (Mckee et al., 1993), as it provided a higher number of images per field to each class, while reducing the number of classes and thus simplifying data analysis.

Subsequently, the images acquired during each one of the periods, were classified accordingly, as occurring during dry, normal or wet weather conditions. The classification resulted in fifteen, partly overlapping, climatic periods for each field, given by the combination of five (three shorter and two longer) periods of the year and three SPEI classes.

2.1.3. Spatial clustering of agricultural fields

Prior to calculating clustering, temporal comparability of NDVI data from the same field was enhanced by applying a Z-score spatial normalization. This returned images with the same standard deviation and zero mean (Pascucci et al., 2018).

To identify spatial patterns of plants vigor, we applied a k-means (k = 3) clustering to each time series of images obtained after climatic classification, hereafter called *climate-based* clustering. As a comparison, we also applied the k-means to the whole set of images available for each field, without climatic distinction, hereafter called *bulk* clustering. This produced up to sixteen clustering results for each field, depending on the availability of images for each combination period and climatic class. Clustering results were represented as maps of the fields, split into three classes. For visualization reasons, an order was given to the classes, so that zones with a high/intermediate/low median NDVI would be represented by the same color in all maps resulting from clustering.

2.2. Comparison of clustering spatial patterns to field soil maps

To test whether the spatial patterns identified by the different clustering procedures corresponded to underlying soil properties, we derived soil property maps for silt, clay, soil organic matter (SOM) and available soil water (ASW) for each field. These were then classified into three classes and compared with the results of the clustering previously computed from the NDVI time-series images.

SOM and soil texture maps (clay, sand and silt) were generated from soil sampling data, spatial covariates and spatial predictions. A total of 5–24 soil samples were collected in each field in 2019, each being a composite of three or more samples collected within a 2 m radius of a georeferenced sampling point, using an Ejkelkamp auger in the 0–40 cm depth layer. Samples were analyzed in the laboratory to determine clay, silt and sand using the pipette method according to the USDA textural thresholds, and SOM content using the Walkley-Black method.

Spatial covariates included apparent electrical resistivity maps, Digital Elevation Models (DEMs) and the Bare Soil Index (BSI) (Chen et al., 2004; Mzid et al., 2021), the latter obtained from Sentinel-2 satellite images (Table 1). Resistivity maps were obtained for some of the fields by means of an Electro Magnetic Agro-Scanner (EMAS) system (SOING, Livorno, Italy) (McBratney et al., 2005). This is an electromagnetic induction (EMI) sensor, that provided measurements of the apparent electrical resistivity for three soil depth layers. In the present work, the 0–50 cm layer was used. The DEMs (spatial resolution of 1 m) were obtained from LiDAR data and were downloaded from the geoportal of the Italian Ministry of the Environment and Territory (http://www.pcn.minambiente.it/viewer/). The use of DEM as covariate was limited to non-flat fields. Additionally, the BSI was used as covariate. We calculated the BSI from Sentinel-2 imagery, acquired after the harvest and before the following crop's emergence on each field, so

that only bare soil or crop residues were visible. The covariates used for each field are reported in Table 1. Soil prediction models were built to spatialize sand, silt, clay and organic matter sampling points to the whole fields, by exploiting the above-mentioned covariates. Either Multiple Linear Regression (MLR), Regression Kriging (RK) or Block Kriging (BK) (Malone et al., 2017) were used, depending on the available covariates and the number of soil samples (Table 1). BK was used where an intensive soil sampling had been carried out (such as for the fields in the Maccarese area), using a block size resolution of 10 m. MLR and RK were employed for the fusion of remotely sensed and geo-physical data, the first when only a few soil samples were available (N = 5), while the second for fields with a higher number of samples. Model performances were evaluated by means of cross-validation, assessing the coefficient of determination (R^2) and Root Mean Squared Error (RMSE).

Soil water content at permanent wilting point and at field capacity were estimated from soil properties (silt, clay and organic matter) using pedotransfer functions (Saxton and Rawls, 2006). Their difference provided an estimate of the maximum plant available soil water (ASW).

2.2.1. Assessment of the matching of NDVI clusterings with soil maps

Soil properties maps were first classified by means of k-means clustering (k = 3). The soil classes resulting from clustering soil maps were ordered to represent high/intermediate/low soil property values with the same colors, as for the zones identified by the NDVI clustering (Section 2.1.3).

The agreement between the clustering on soil properties maps and on the normalized NDVI time series was evaluated by means of the Adjusted Rand Index (ARI) (Albatineh et al., 2006; Heil et al., 2019). ARI indicates the degree by which couples of pixels, either belonging to the same or to different classes in a map, do the same in the companion map. It takes a value near to zero for correspondence of random sets, and a value of one for full correspondence, with increasing values for increasing similarity between the maps.

The clustering procedure is meant to identify soil property classes, especially when soil presents relatively large variability. In our case, clustering performance varied considerably both across and within clustering parameter sets. To elucidate whether clustering performances in matching soil property maps were affected by soil heterogeneity of the different soil properties, we assessed the linear relationships

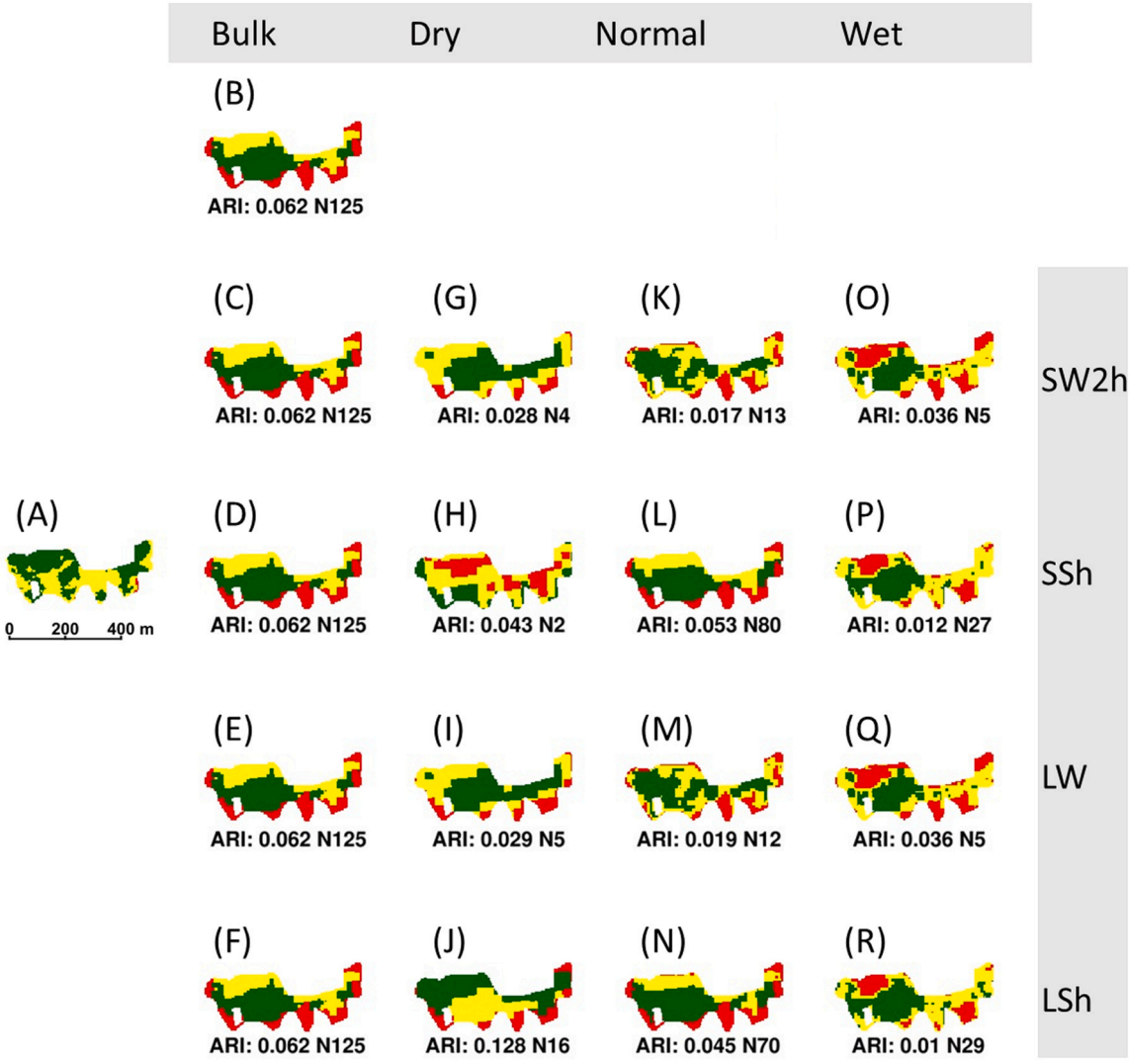

Fig. 3. Maps resulting from applying k-means clustering (k = 3) to the different NDVI time-series (bulk and classified according to the climate) for an arable crop field (Bertoldo). The Available Soil Water (ASW) map obtained from clustering of ground data is reported for comparison (A). Maps (B) to (F) were obtained by bulk clustering (all NDVI images); maps (G) to (J) from clustering only images from dry years; maps from (K) to (N) only from images from normal weather years; maps (O) to (R) from images from wet years. Rows indicate clustering time span: see text for code meaning. Below each map are reported the Adjusted Rand Index (ARI) and the number of images used for the clustering.

between heterogeneity (expressed by the coefficient of variation) in soil properties and the correspondence (in terms of ARI) between the NDVI series and each soil property map. A comparison of the slopes of linear relationships was then used to assess the different sensitivity (paired t-tests) to soil heterogeneity of the best performing climatic clustering in respect to the bulk clustering.

3. Results

The number of NDVI images selected according to the criteria defined above and thus available for the following clustering procedure, strongly depended on the occurrence of rather specific weather conditions (Supplementary Table 1). Especially for the short-early season (e.g. SPEI SW1h) sometimes no images were available, because particularly dry and wet years according to the SPEI classification did not occur, and thus clustering could not be performed. This occurred for example for the Bertoldo field (Fig. 3) for which no results were available for SW1h. This is likely due to the cloudy sky and low vegetation vigor in winter, since when field images had less than 70% of pixels with NDVI higher than 0.7 they were discarded, in addition to fewer years with particularly dry or wet seasons in respect to normal ones. From Fig. 3 it appears clearly that bulk clustering (Figs. 3b to 3f) and many of the climate-based clustering results did not match what was observed from ground soil mapping data (Fig. 3a), i.e. identified a reversed pattern in terms of high or medium soil AWC. However, the clustering carried out considering only NDVI images occurring during dry seasons, covering the period from January to May (LSh), was able to identify the correct AWC pattern (Fig. 3j) and had an ARI value of 0.128, the highest among all clustering results.

In general, the Adjusted Rand Index (ARI), used for comparing clustering results to soil maps, showed a remarkable variability across agricultural fields for all soil properties. A higher median ARI value was often associated with a larger variability (Fig. 4). The mean ARI for arable fields covered by herbaceous species, varied depending on the soil property considered, with the highest ARI values for sand (min = 0.027, max = 0.165, mean = 0.093), then clay (0.024, 0.116, 0.087), ASW (0.022, 0.120, 0.080), organic matter (0.036, 0.112, 0.076) and silt (0.030, 0.128, 0.071) (Fig. 4, Table 3). For clay, bulk clustering obtained better results than climate-based clustering, for all phenological phases. However, for all other soil properties, climate-based clustering was consistently better than bulk clustering, when performed over long spring periods (LSh). Bulk clustering obtained higher ARI values than most climate-based early season clustering (SW1h, SW2h and LW). The clustering obtained over the short spring period (SSh), i.e. from March to May, had ARI values similar or lower than the bulk clustering for dry seasons, and generally similar or even higher for wet seasons, except for ASW. The clustering that best matched soil maps, both for sand, silt and ASW, was based on images acquired during the long spings (LSh) in dry years (Table 3). Considering the match with clay maps, the wet LW clustering gave a slightly better score than the LSh (0.117 vs 0.112), but

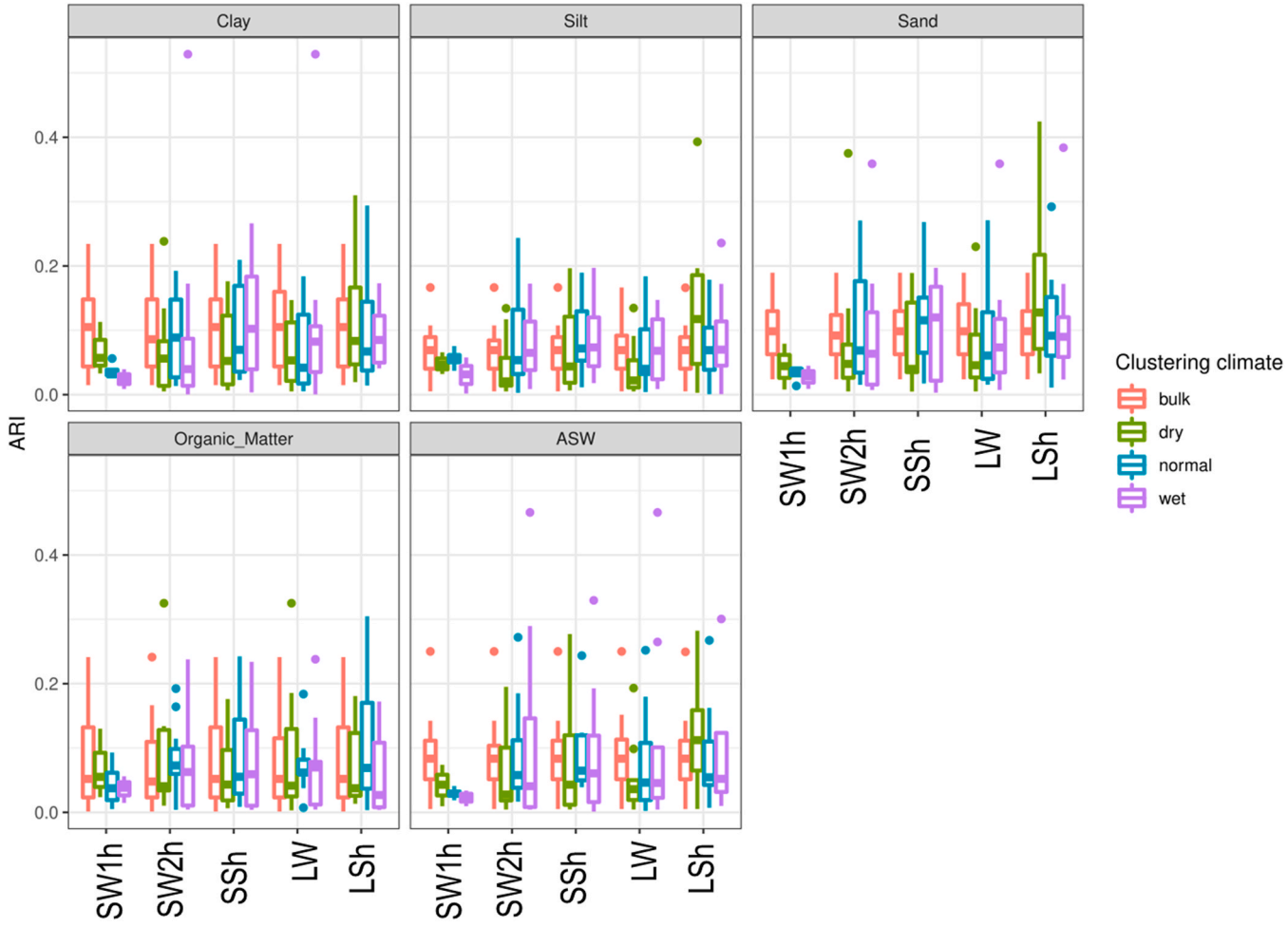

Fig. 4. Boxplots of the Adjusted Rand Index (ARI) obtained from the comparison among clustering results and ground truth maps of soil properties. Clustering: "bulk" includes all images, independent of the climate; "dry" / "normal" / "wet" includes only images from years classified accordingly, based on the SPEI index. Individual dots are values beyond 1.5 * IQR (interquartile range). The time periods over which images were extracted are: November to January (SW1h), January to March (SW2h), March to May (SSh), November to March (LW) and January to May (LSh).

Table 3
Adjusted Rand Index (ARI) for the agreement between clustering applied to the NDVI time-series and each soil property map for arable crops. The three clustering parameter sets with highest mean ARI are reported, with associated standard deviation and coefficient of variation (CV).

Soil property	Clustering time span	Clustering climate	ARI		
			mean	standard deviation	coefficient of variation
Sand	LSh	dry	0.165	0.133	0.80
	SSh	normal	0.117	0.068	0.58
	LSh	wet	0.116	0.104	0.89
Silt	LSh	dry	0.128	0.111	0.86
	SSh	wet	0.090	0.065	0.72
	LSh	wet	0.090	0.07	0.78
ASW	LSh	dry	0.120	0.089	0.74
	SW2h	wet	0.119	0.169	1.42
	LW	wet	0.113	0.154	1.36
Clay	LW	wet	0.117	0.162	1.39
	LSh	dry	0.112	0.09	0.8
	SSh	wet	0.112	0.09	0.8
Soil Organic Matter	LSh	normal	0.112	0.098	0.87
	LW	dry	0.092	0.102	1.11
	SSh	normal	0.091	0.081	0.89

with much larger variability (respectively CV = 139% vs 80%) (Table 3). Considering organic matter, its best mean score was obtained once again by a long spring clustering (LSh), but under normal climatic conditions rather than dry ones.

The performance in matching soil maps seemed to be at least partly field dependent, with some fields obtaining a relatively short range of ARI values (e.g. Biavati), no matter the clustering parameter set, while others (e.g. Tamburini) achieving a wider distribution. The variability across fields is shown in Fig. 5 for ASW. Climate-based clustering results obtained higher ARI values, i.e. matched better the ground truth ASW maps, than bulk clustering for all fields. Despite the highest mean values presented by the LSh dry clustering, other late season climate clusterings often obtained the best matches with soil maps in individual fields (Fig. 5). The number of images available for each clustering spanned across two orders of magnitude, depending on the clustering parameters set. Nonetheless, clustering parameter sets with the highest number of images never obtained the best match, which was generally obtained by clusterings based on 25 images ore less (Fig. 5).

Concerning tree crop fields, clustering results matched soil properties with ARI values similar to herbaceous fields, but with highest mean values obtained for clay (min = 0.087, max = 0.361, mean = 0.204), then for ASW (0.039, 0.410, 0.190), silt (0.030, 0.210, 0.135), sand (0.024, 0.156, 0.067) and finally organic matter (0.008, 0.125, 0.056) (Fig. 6).

The limited number of years for which NDVI time-series were available for tree crops made it impossible to obtain most of the dry climate clusterings (only except the SSt), and limited the variety of wet climate clusterings. However, except for organic matter for the olive field (Zuccari), climate-based clustering always provided better correspondence with soil maps than the bulk clustering. In particular, the period from April to August (LSt2), for wet years, gave the best results for the vineyard (Bennicelli) for all soil properties except for silt. The same type of climate-based clustering was not available for the olive field (Zuccari), for which the shorter summer period, i.e. June to August (SSmt), for normal years provided the best correspondence, except for organic matter. Similarly to what observed for fields covered by arable crops, also in this case the highest scores were more related to the specific soil parameter than to a higher number of images.

Considering all the fields, both covered by arable and tree crops, in general the climate-based clustering that best matched soil maps was the one based on the weather pattern occurring in late season, for longer periods (LSh and LSt2). In the case of herbaceous species, the relatively higher number of available fields allows to investigate whether the mean best match, obtained by the LSh clustering in dry years, is due to some particularly high scoring field, or to a more general good agreement with soil maps. For this reason, results of the dry years LSh climate-based clustering are presented field-wise, in comparison to the bulk

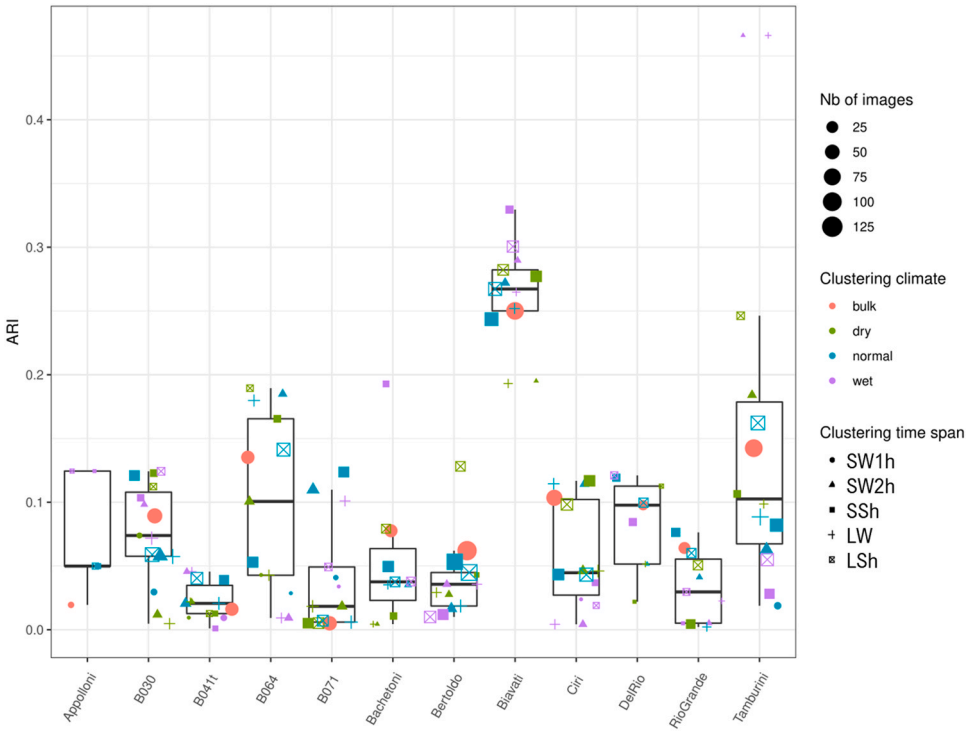

Fig. 5. Boxplots of Adjusted Rand Index (ARI) for available soil water (ASW) in bulk and climate-based clustering, for each field covered by herbaceous crops. Dots indicate individual ARI values, symbols indicate the clustering time span, while their size is proportional to the number of images that contributed to the individual clustering.

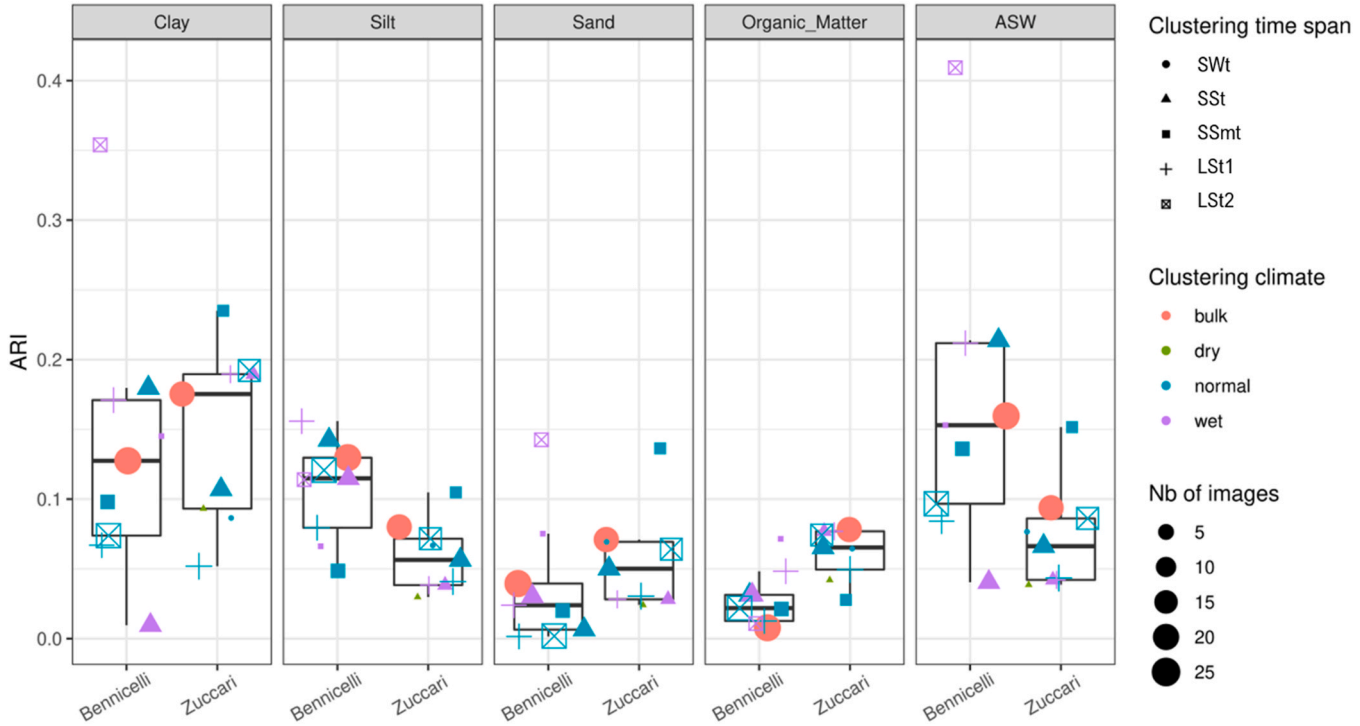

Fig. 6. Boxplots of Adjusted Rand Index (ARI) for the fields covered by tree crops, for bulk and climate-based clustering. Dots indicate individual ARI values, symbols indicate the clustering time span, while their size is proportional to the number of images that contributed to the individual clustering.

clustering in Fig. 7. ARI of the dry LSh clustering was systematically higher in respect to the bulk clustering for almost all fields for silt, sand and ASW, whereas it provided results similar or alternately better and worse, depending on the field, for clay and organic matter (Fig. 7). As mentioned above, the best results for soil organic matter were obtained by the long spring clustering (LSh), based on normal climate years, which outpaced the bulk clustering on nine fields out of twelve (Fig. 8).

As the correspondence between clusterings and soil maps varied greatly, depending on the field, even for the best scoring seasonal clustering, we questioned whether some characteristics of the fields may be related to better or worse results. Inspection of the correlation between the clustering scores (ARI) obtained for the different soil properties and some soil properties statistics (mean, skewness and coefficient of variation), indicated that the coefficient of variation (CV) was the metric with the highest correlation values with ARI. In particular, ARI values obtained by the LSh clustering for dry years presented a higher number of significant correlations ($n = 5$) than the bulk clustering ($n = 3$).

Besides any specific significant threshold, the slopes of the relationships between the variability in soil properties and ARI (Fig. 9), were significantly higher ($P < 0.05$) for the LSh clustering in dry years (0.468 ± 0.167) than for the bulk clustering (0.113 ± 0.270), indicating a higher sensitivity of the climate-based clustering in matching soil properties, when these are highly variable. In greater detail, while relationships with similar slopes were obtained for clay and ASW, a higher sensitivity was present for silt, sand and organic matter.

Finally, considering the processing time after image download, the time required by a laptop computer (Asus N552VX, CPU: Intel Core i7 6700HQ, RAM: 8 Gb DDR4, HD: SSD) to process images for a relatively large field (B030: 750 ×270 m) consisted on 3646 s for image preprocessing, 172 s for climatic image classification and 409 s for performing all clusterings (bulk and climate-based), and saving outputs.

4. Discussion

The need to improve the efficiency of agricultural management, by adoption of site-specific agronomic practices, has been constrained by the lack of information on the actual soil properties patterns occurring within agricultural fields. A very widespread precision agriculture approach is that of subdividing fields into smaller sub-units, sometimes called uniform management zones (MZ), which are considered to have a rather more homogeneous response to environmental and agronomic factors, as compared to the whole field (Nawar et al., 2017). Naturally, this more homogeneous response of MZs depends largely on the homogeneity of the underlying soil properties. Several methods based on the analysis of multi-temporal remote sensing imagery have been proposed for zoning agricultural fields and identifying the patterns of soil spatial variability. These include those based solely on remote sensing data (e.g. Blasch et al., 2015; Pascucci et al., 2018), as well as those including other co-variates such as topography, proximal sensing geophysical data, soil sampling and yield maps (e.g. Khosla et al., 2008; Miao et al., 2018; Scudiero et al., 2018), or plant structural properties in interaction with climate, such as LAI and evapotranspiration (Ohana--Levi et al., 2021; Shuai and Basso, 2022). The results of different zoning procedures obtained for one field from different approaches tend to differ, depending not only on the data source, but also on the clustering method used (Pascucci et al., 2018). From the perspective of the user, it is important to know how the different sub-areas of the field, or MZs, will respond in different years characterized by contrasting weather patterns. Maestrini and Basso (2018) observed that temporally stable or unstable yielding zones of agricultural fields, across years, are characterized by different topographic wetness indices. Areas characterized by a low topographic wetness index (proxy for areas with probability of lower water content) always perform poorly (low and stable yield), whereas mid-high wetness index areas have high and stable yield. Unstable areas are those with a very high wetness index, corresponding to depression areas, waterlogged and poorly yielding in wet years, but high yielding in dry years. These authors emphasize the need for a different management approach, respectively strategic or tactical, for stable and unstable yielding areas of a field.

To our knowledge, the present study is the first attempt of exploiting the use of climatic indices, together with multi-spectral indices of crop

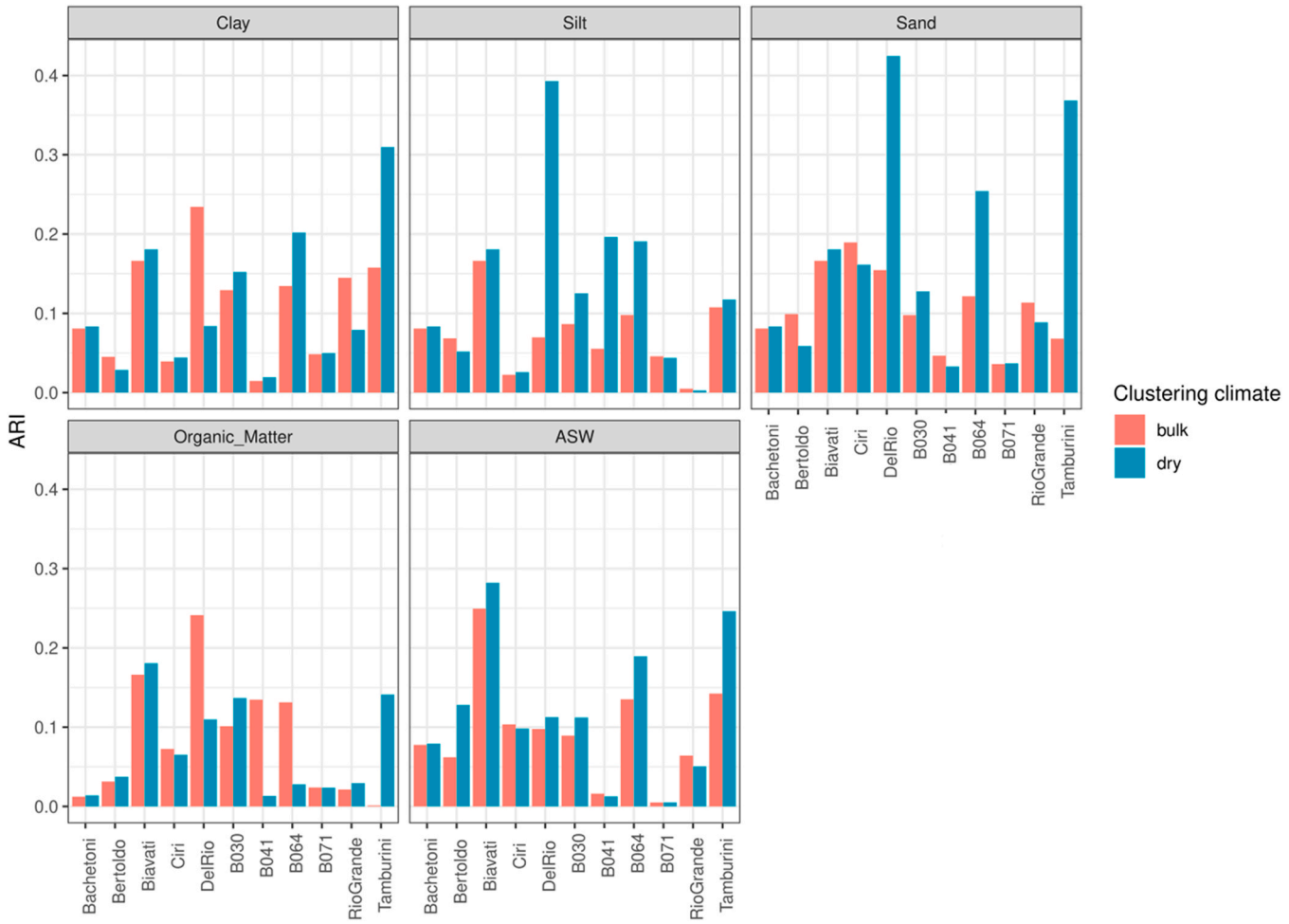

Fig. 7. ARI obtained by the long spring (LSh) clustering in dry years and bulk clusterings for each field and soil properties maps, for fields covered by herbaceous species. Appolloni field was removed as a LSh clustering in dry years was not available for it.

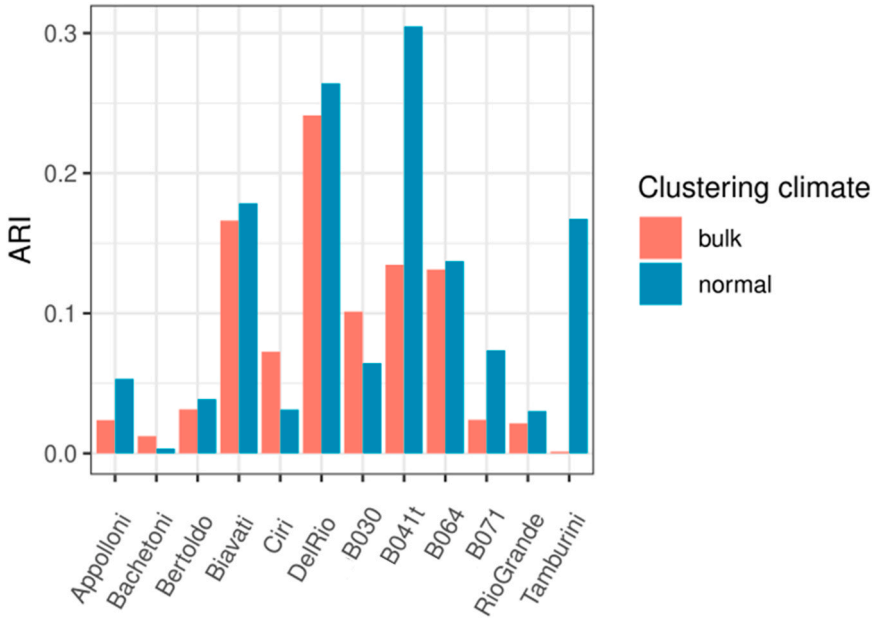

Fig. 8. ARI obtained by the comparison of climate-based and bulk clustering, for the time span from January to May (LSh), with the soil organic matter maps, for each field with herbaceous species.

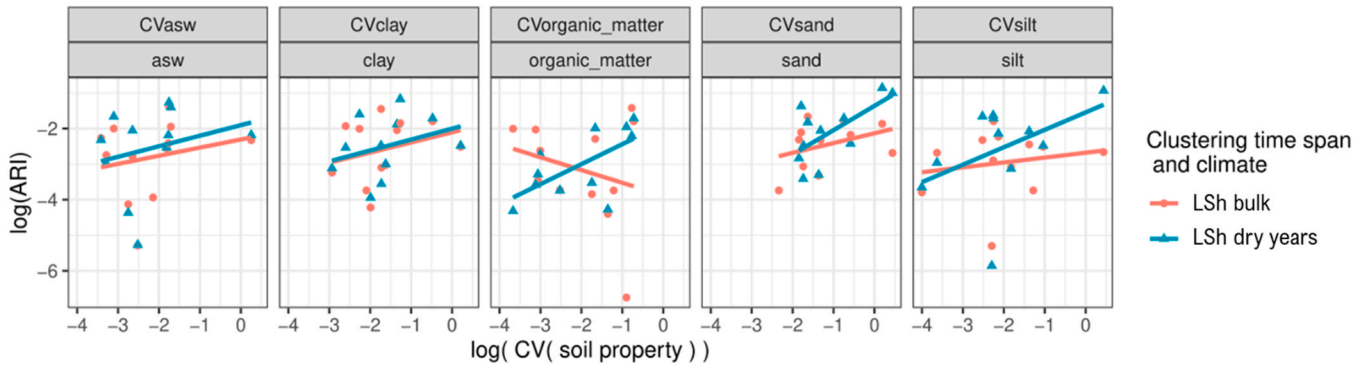

Fig. 9. Relationships between ARI and the coefficient of variation of soil properties for the bulk clustering and the climate-based clustering including images from January to May (LSh) in dry years. Values are expressed in logarithmic terms to obtain normal distributions.

vigor, to improve the identification of soil property zones. The main hypothesis of the proposed method is that a specific seasonal weather pattern may exacerbate the differences in suitability for crop growth, and thus plant response, resulting from the different soil conditions present within a field. The higher correspondence between multiple climate-based clusterings and soil maps, in respect to the bulk clustering, confirms the soundness of this hypothesis.

The assessment of the climate-based clusterings obtained by different parameters sets showed that several of them outpaced the conventional bulk clustering (Table 3), in both herbaceous and tree crop fields (Figs. 4, 5, 6). In addition, the deeper analysis on the best scoring climate-based clustering on herbaceous fields (Figs. 7, 8) showed that this performed consistently better than the bulk clustering in almost all analyzed fields, especially for silt, sand and ASW. In short, our results demonstrate that climatic criteria are appropriate to improve filtering of RS imagery for the identification of MZ.

We highlighted that the clustering parameter set had a far larger influence in matching soil properties, than the number of images contributing to the clustering (Figs. 5, 6). Improvements in matching soil properties cannot come mostly from increasing the size of the acquired RS dataset, but should be based on a greater attention to the climate-crop-field interactions, leading a thoughtful process of image selection.

The aim of a clustering is to find the differences present in a field, especially where these are major. In this respect, it is important to stress that the climate-based LSh clustering for dry years seemed to be more sensitive to soil variability than the bulk clustering (Fig. 9), increasing in accuracy where most needed, namely in fields with more heterogeneous soil properties. This property is fundamental, as it implies a higher probability to correctly identify soil classes where higher heterogeneity is present, and therefore distinction is needed.

The most effective climate for selecting appropriate imagery was the dry one for herbaceous species. In a dry climate, the evapotranspiration demand is high in respect to precipitation, likely determining water limiting conditions for the plants. Considering the increases in the evapotranspiration demand and relative imbalance with precipitations due to global warming (Lionello and Scarascia, 2018) it is likely that image selections driven by climate criteria may become increasingly effective in identifying homogeneous soil zones over time.

As a second hypothesis, a crop was thought to immediately respond (as portrayed by instantaneous NDVI snapshots) to a given climate, thus permitting to split the images acquired during a single crop growth cycle into multiple groups, each one reflecting plant responses possibly to contrasting climates. As an example, a field characterized by sandy and clay soil zones may show high differences in plants growth across zones after a water deficient early season period, but may recover after a subsequent wet late season. Our results do not contradict the validity of this hypothesis but, by obtaining better results for longer climatic periods, they suggest that the cumulative effects of climate on plant development over long periods are more effective for identifying subtending soil zones than shorter term effects of possibly more extreme climates.

The effects of a given weather on plants may be somewhat delayed, especially for deep rooted plants such as perennials, and for drought tolerant species (such as grapevine and olive). An assessment of this delay would need to be made before introducing such a feature into the selection of imagery.

The presented method classifies climate and images of the crop based on fixed periods: this implies cutting off images outside pre-defined dates. Considering that the dates of a plant growth cycle vary from year to year, and that clusterings based on long climatic periods performed best, we suggest that future development of a climate-based clustering should include image selection based on dynamic dates, following phenological development. This would allow to more closely follow the crop growth cycle, and the cumulative effects of weather on plant development.

The limited number of available fields may have restricted the power of our analysis, indeed our tests were run for a total of 14 fields, 12 of which with herbaceous and two with tree crops. Despite the first results for the case of tree crops are promising, a larger number of sample fields occupied by the same tree species for a long time would likely allow to draw more solid conclusions on best clustering parameters sets for tree crops as well. In addition, any pixel from the considered satellites may contain signals from multiple rows and inter-rows, so that the proposed method cannot discriminate between adjacent plants. As such, it should be considered that, when applied to tree crops, this method assesses an average vegetation vigor, integrating both the target tree species and herbaceous species possibly present around the trees, both responding to the soil conditions. In addition, the derived NDVI values may underestimate the differences between high and low vigor plants when images are acquired soon after mowing or tree summer pruning. Further, also in the case of herbaceous fields, soil maps based on a higher number of sampling points, would lower the effects of random errors, thus increasing the power of the analysis.

5. Conclusions

This is, to the authors knowledge, the first soil zoning study based on clustering remote sensing imagery, in which the identification of management zones in agricultural fields is improved by integrating a climatic index among the image selection criteria. Our analysis on NDVI time-series revealed that, both the climatic conditions occurring in the field during image acquisition and the considered crop growth period play a major role in the ability of a clustering algorithm to identify soil property zones. In particular, using time-series from long, dry and wet, late growth season periods significantly increased the ability to identify homogeneous soil zones, in respect to the absence of climatic criteria, both in herbaceous and tree crop fields, and consistently across most of them. Further, the method showed to be significantly more sensitive to

soil heterogeneity, in respect to the reference bulk clustering, thus better identifying management zones in fields where a proper soil zoning is most needed. This study highlights that considering the climate-crop-field interactions is crucial to increase the accuracy in the identification of management zones in agricultural fields. The proposed method is likely to improve its performance over time, when applied to the increasingly water stressed crops undergoing the exacerbated climatic extremes imposed by global warming.

CRediT authorship contribution statement

Francesco Reyes: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft; Writing – review & editing. **Raffaele Casa:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft. **Massimo Tolomio:** Data curation, Formal analysis, Writing – original draft. **Michele Dalponte:** Data curation, Formal analysis, Writing – original draft. **Nada Mzid:** Data curation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The authors do not have permission to share data.

Acknowledgments

The authors acknowledge the SMARTAGRI and VISTA projects, P.S. R. Mis. 16.2 of the Umbria Region, for contributing to fund this research.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eja.2023.126930.

References

Albatineh, A.N., Niewiadomska-Bugaj, M., Mihalko, D., 2006. On similarity indices and correction for chance agreement. J. Classif. 23, 301–313. https://doi.org/10.1007/S00357-006-0017-Z.

Albornoz, E.M., Kemerer, A.C., Galarza, R., Mastaglia, N., Melchiori, R., Martínez, C.E., 2018. Development and evaluation of an automatic software for management zone delineation. Precis. Agric. 19, 463–476. https://doi.org/10.1007/s11119-017-9530-9.

ARSIAL -L'Agenzia Regionale per lo Sviluppo e l'Innovazione dell'Agricoltura del Lazio [WWW Document], n.d. URL ⟨https://www.arsial.it/⟩ (accessed 12.23.21).

Basso, B., Ritchie, J.T., Cammarano, D., Sartori, L., 2011. A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. Eur. J. Agron. 35, 215–222. https://doi.org/10.1016/j.eja.2011.06.004.

Beguería, S., Vicente-Serrano, S.M., 2018. Index: SPEI, The Standardised Precipitation-Evapotranspiration Index [WWW Document]. URL ⟨https://spei.csic.es/⟩ (accessed 12.22.21).

Blasch, G., Spengler, D., Hohmann, C., Neumann, C., Itzerott, S., Kaufmann, H., 2015. Multitemporal soil pattern analysis with multispectral remote sensing data at the field-scale. Comput. Electron. Agric. 113, 1–13. https://doi.org/10.1016/J.COMPAG.2015.01.012.

Breunig, F.M., Galvão, L.S., Dalagnol, R., Dauve, C.E., Parraga, A., Santi, A.L., della Flora, D.P., Chen, S., 2020. Delineation of management zones in agricultural fields using cover–crop biomass estimates from PlanetScope data. Int. J. Appl. Earth Obs. Geoinf. 85, 102004 https://doi.org/10.1016/J.JAG.2019.102004.

Castrignanò, A., Buttafuoco, G., Quarto, R., Vitti, C., Langella, G., Terribile, F., Venezia, A., 2017. A combined approach of sensor data fusion and multivariate geostatistics for delineation of homogeneous zones in an agricultural field. Sensors 17, 2794. https://doi.org/10.3390/s17122794.

Castrignanò, A., Buttafuoco, G., Quarto, R., Parisi, D., Viscarra Rossel, R.A., Terribile, F., Langella, G., Venezia, A., 2018. A geostatistical sensor data fusion approach for delineating homogeneous management zones in Precision Agriculture. Catena 167, 293–304. https://doi.org/10.1016/j.catena.2018.05.011.

Chen, C.C., McCarl, B.A., Schimmelpfennig, D.E., 2004. Yield variability as influenced by climate: a statistical investigation. Clim. Change 66, 239–261. https://doi.org/10.1023/B:CLIM.0000043159.33816.e5.

Danlu, G., Seth, W., Tim, P., 2020. Evapotranspiration: Modelling Actual, Potential and Reference Crop Evapotranspiration. R package version 1.15. [WWW Document]. URL ⟨https://cran.r-project.org/package=Evapotranspiration⟩.

Delerce, S., Dorado, H., Grillon, A., Rebolledo, M.C., Prager, S.D., Patiño, V.H., Varón, G. G., Jiménez, D., 2016. Assessing weather-yield relationships in rice at local scale using data mining approaches. PLoS ONE 11. https://doi.org/10.1371/journal.pone.0161620.

Dhawale, N.M., Adamchuk, V.I., Prasher, S.O., Dutilleul, P.R.L., Ferguson, R.B., 2014. Spatially constrained geospatial data clustering for multilayer sensor-based measurements. ISPRS - Int. Arch. Photogramm., Remote Sens. Spat. Inf. Sci. XL–2 187–190. https://doi.org/10.5194/isprsarchives-XL-2-187-2014.

Diker, K., Heermann, D.F., Brodahl, M.K., 2004. Frequency analysis of yield for delineating yield response zones. Precis. Agric. 5, 435–444. https://doi.org/10.1007/S11119-004-5318-9.

Fontanet, M., Scudiero, E., Skaggs, T.H., Fernàndez-Garcia, D., Ferrer, F., Rodrigo, G., Bellvert, J., 2020. Dynamic management zones for irrigation scheduling. Agric. Water Manag. 238. https://doi.org/10.1016/j.agwat.2020.106207.

Franzen, D., Nanna, T., 2002. Management zone delineation methods. In: R., P.C., R., R. H., L., W.E. (Eds.), Sixth Int. Conf. on Precision Agriculture. Minneapolis.

Gavioli, A., de Souza, E.G., Bazzi, C.L., Schenatto, K., Betzek, N.M., 2019. Identification of management zones in precision agriculture: an evaluation of alternative cluster analysis methods. Biosyst. Eng. 181, 86–102. https://doi.org/10.1016/j.biosystemseng.2019.02.019.

Ge, Y., Thomasson, J.A., Sui, R., 2011. Remote sensing of soil properties in precision agriculture: a review. Front. Earth Sci. https://doi.org/10.1007/s11707-011-0175-0.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/J.RSE.2017.06.031.

Heil, J., Häring, V., Marschner, B., Stumpe, B., 2019. Advantages of fuzzy k-means over k-means clustering in the classification of diffuse reflectance soil spectra: a case study with West African soils. Geoderma 337, 11–21. https://doi.org/10.1016/J.GEODERMA.2018.09.004.

Johnson, C.K., Mortensen, D.A., Wienhold, B.J., Shanahan, J.F., Doran, J.W., 2003. Site-specific management zones based on soil electrical conductivity in a semiarid cropping system. Agron. J. 95, 303–315. https://doi.org/10.2134/AGRONJ2003.0303.

Khosla, R., Inman, D., Westfall, D.G., Reich, R.M., Frasier, M., Mzuku, M., Koch, B., Hornung, A., 2008. A synthesis of multi-disciplinary research in precision agriculture: Site-specific management zones in the semi-arid western Great Plains of the USA. Precis. Agric. 9, 85–100. https://doi.org/10.1007/S11119-008-9057-1.

Lionello, P., Scarascia, L., 2018. The relation between climate change in the Mediterranean region and global warming. Reg. Environ. Change 2018 18:5 18, 1481–1493. https://doi.org/10.1007/S10113-018-1290-1.

Maestrini, B., Basso, B., 2018. Predicting spatial patterns of within-field crop yield variability. Field Crops Res. 219, 106–112. https://doi.org/10.1016/j.fcr.2018.01.028.

Malone, B.P., Minasny, B., McBratney, A.B., 2017. Using R for digital soil mapping, Vol. 35. Springer International Publishing, Cham, Switzerland.

Marino, S., Alvino, A., 2018. Detection of homogeneous wheat areas using multi-temporal UAS images and ground truth data analyzed by cluster analysis. Eur. J. Remote Sens. 51, 266–275. https://doi.org/10.1080/22797254.2017.1422280.

McBratney, A.B., Minasny, B., Whelan, B.M., 2005. Obtaining 'useful'high-resolution soil data from proximally-sensed electrical conductivity/resistivity (PSEC/R) surveys. Precis. Agric. 5, 503–510.

Mckee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. Eighth Conf. Appl. Climatol. 179–183.

Miao, Y., Mulla, D.J., Robert, P.C., 2018. An integrated approach to site-specific management zone delineation. Front. Agric. Sci. Eng. 5, 432–441. https://doi.org/10.15302/J-FASE-2018230.

Moore, F.C., Lobell, D.B., 2014. Adaptation potential of European agriculture in response to climate change. Nat. Clim. Change 4, 610–614. https://doi.org/10.1038/nclimate2228.

Mulla, D.J., 2013. Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. Biosyst. Eng. 114, 358–371. https://doi.org/10.1016/j.biosystemseng.2012.08.009.

Mzid, N., Pignatti, S., Huang, W., Casa, R., 2021. An analysis of bare soil occurrence in arable croplands for remote sensing topsoil applications. Remote Sens. 13, 1–24. https://doi.org/10.3390/RS13030474.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., Mouazen, A.M., 2017. Delineation of soil management zones for variable-rate fertilization: a review. Adv. Agron. 143, 175–245. https://doi.org/10.1016/BS.AGRON.2017.01.003.

Ohana-Levi, N., Bahat, I., Peeters, A., Shtein, A., Netzer, Y., Cohen, Y., Ben-Gal, A., 2019. A weighted multivariate spatial clustering model to determine irrigation management zones. Comput. Electron. Agric. 162, 719–731. https://doi.org/10.1016/j.compag.2019.05.012.

Ohana-Levi, N., Gao, F., Knipper, K., Kustas, W.P., Anderson, M.C., del Mar Alsina, M., Sanchez, L.A., Karnieli, A., 2021. Time-series clustering of remote sensing retrievals for defining management zones in a vineyard. Irrig. Sci. https://doi.org/10.1007/S00271-021-00752-0.

Ortega, R.A., Santibáñez, O.A., 2007. Determination of management zones in corn (Zea mays L.) based on soil fertility. Comput. Electron. Agric. 58, 49–59. https://doi.org/10.1016/J.COMPAG.2006.12.011.

Osborne, T.M., Wheeler, T.R., 2013. Evidence for a climate signal in trends of global crop yield variability over the past 50 years. Environ. Res. Lett. 8, 024001 https://doi.org/10.1088/1748-9326/8/2/024001.

Pascucci, S., Carfora, M.F., Palombo, A., Pignatti, S., Casa, R., Pepe, M., Castaldi, F., 2018. A comparison between standard and functional clustering methodologies: application to agricultural fields for yield pattern assessment. Remote Sens. 10. https://doi.org/10.3390/rs10040585.

Powell, J.P., Reinhard, S., 2015. Measuring the effects of extreme weather events on yields. Weather Clim. Extrem. 12, 69–79. https://doi.org/10.1016/j.wace.2016.02.003.

R Core Team (2020). — European Environment Agency [WWW Document], n.d. URL ⟨https://www.eea.europa.eu/data-and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-2006⟩ (accessed 2.21.22).

Ray, D.K., Gerber, J.S., Macdonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6, 1–9. https://doi.org/10.1038/ncomms6989.

Saifuzzaman, M., Adamchuk, V., Buelvas, R., Biswas, A., Prasher, S., Rabe, N., Aspinall, D., Ji, W., 2019. Clustering tools for integration of satellite remote sensing imagery and proximal soil sensing data. Remote Sens. 11, 1036. https://doi.org/10.3390/rs11091036.

Saxton, K.E., Rawls, W.J., 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci. Soc. Am. J. 70, 1569–1578. https://doi.org/10.2136/SSSAJ2005.0117.

Schepers, A.R., Shanahan, J.F., Liebig, M.A., Schepers, J.S., Johnson, S.H., Luchiari, A., 2004. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agron. J. 96, 195–203. https://doi.org/10.2134/AGRONJ2004.1950.

Scudiero, E., Teatini, P., Manoli, G., Braga, F., Skaggs, T.H., Morari, F., 2018. Workflow to establish time-specific zones in precision agriculture by spatiotemporal integration of plant and soil sensing data. Agronomy 2018 Vol. 8, 253. https://doi.org/10.3390/AGRONOMY8110253.

Shuai, G., Basso, B., 2022. Subfield maize yield prediction improves when in-season crop water deficit is included in remote sensing imagery-based models. Remote Sens. Environ. 272, 112938 https://doi.org/10.1016/J.RSE.2022.112938.

Vicente-Serrano, S.M., Beguería, S., 2018. SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index. R package version 1.7.2. [WWW Document]. URL ⟨http://spei.csic.es⟩.

Wang, H., Vicente-serrano, S.M., Tao, F., Zhang, X., Wang, P., Zhang, C., Chen, Y., Zhu, D., Kenawy, A. el, 2016. Monitoring winter wheat drought threat in Northern China using multiple climate-based drought indices and soil moisture during 2000–2013. Agric. For. Meteorol. 228–229, 1–12. https://doi.org/10.1016/j.agrformet.2016.06.004.

Zhang, N., Wang, M., Wang, N., 2002. Precision agriculture—a worldwide overview. Comput. Electron. Agric. 36, 113–132. https://doi.org/10.1016/S0168-1699(02)00096-0.